

SEQ ID NO:1

Tc	NRKSVCPKQKFFFSAPFFFFFCVFPPLISRTGQEKLLFDQKYKLIKGEKKKKKKKNQRANRREHQKREIMRFXNS	75
Tc	FTCIDMHTEGEAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIIEGADLGMYF	151
Tc	MDTGGYLNMCCHNSIAAVTRAIVETGIVSVPAKATNPVWLDTPAGLVRGTAHLQSGTESEVSNASTINVPSTLYQ	225
Tc	QDVVVVLPKPYGEVRVDIAFGGNFTAIVPAEQLGIDISVQNLRLQEAGELLRTEINRSVKVQHPQLPHINTVDC	300
Tc	VEIYGPPTNPEANYKVVVIFGNRQADR	
	<div style="border: 1px solid black; display: inline-block; padding: 2px 10px;">SPCGT</div> GTSAKMATLYAKGQLRIGETFVYESILGSLFQGRV--LGEE	371
Tc	RIPGVKVPVTKDAEEGMLVVTAETGKAFINGFNTMLFDPTDPTKNGFTLKQ*	423

SEQ ID NO:2

Tc	RTGQEKLLFDQKYKIIKGEKKEKKQNPANRREHQCKREIMRFKKS	75
Tc	FTCIDMHTEGEAAARIVTSGLPHIPGSNMAEKKAYLQENMDYLRRGIMLEPRGHDDMFGAFLFDPIIEGADLGWVF	150
Tc	MDTGGYLNMCQHNSIAAVTAAVETGIVSVPAKATNPVVLDTAGLVRGTAHLQSGTISEVSNASIIINVPSFLYQ	225
Tc	QDVVVVLPKPYGEVRVDIAFGGNFFAIVPAEQQLGIDISVQNLSRLQEAGELLRTEINRSVKVQHPQLPHINTVDC	300
Tc	VEIYGPPPTNPEANYKNVVIFGNROADR	
	SPCGT	
Tc	GTSAKMATLYAKGQLRIGETTFVYESILGSLFQGRV--LGEE	371
Tc	RIPGVKVPVTKDAEEGMLVVTAEITGKAFIMGFNTMLFDPTDPFKNGFTLKQ*	423

SEQ ID NO:3

CC MRKSVCPKQKTTFSATPTTTTTCVFPLIS

SEQUENCE ID NO.4

73
150
225
300
375
423

SEQUENCE ID NO:5

Cs -----MKFSKG 6
 Cs IHAIDSHTMGEPTTRIVVGGIPQINGETMAOKKKYLEDNLDYVRTALMHEPRGHNDMFGSIITSSNNKEADFGIIF 31
 Cs MDGGGYLNMCGHGSIGAA TVAVETGMVEMVEPVTNIN--MEAPAGLIKAKVMVEN---EKVKEVSIQNWPSFLYM 151
 Cs EDAKLEVPSLNKTITFDISFGGSFFAIHAKELGWKVETSQVDVLKKLGIEIROLINENIKVQHPELEHIXTVDL 226
 Cs VEIYDEPSNPEATYKNVVIFGQGQVDR SPCGT GTSAKLATLYKKGHLKIDEKEVYESITGTMFKGRV--LEET 297
 Cs KVGEFD-----AIIPETGGAYITGENHEVIDPEDPLKYGFTV*-- 335

SEQ ID NO:6

Pa -----MQR 1
Pa IRIIDSHTGGEPTRLVIGGFPPDLGQGDMAERRRLGERNDANRAACILEPRGSDYLVGALLCAPVCPACAGVIF 79
Pa FNNSGYLGMCGHGTIGLVASLAHLGRIGPGV-----HRIETPVGEVEATLH-----EDGSVSURNVPAYRYR 140
Pa RQVSVEVPGI-GRVSGDIAGGNNWFFLVAGH--GQRLAGDNLDALTANTVAVQQALDD----QDIRGEDGGAICH 219
Pa IELFAD--DPHADSRNFVLCPGKAYDR SPCGT GTSAKLAACLAADGKLLPGQPWRQASVIGSQFEGRYEWLDGQ 279
Pa PGGPIVPTIRGRAHVSAEATLLLADDDPFAWGIRR* ----- 314

SEQ ID NO: 7

Splice leader
acceptor sites

Signal Peptide

Polypyrimidine rich region



CCTTTTCTTTTAAAAACAAAAAAATTCCGGGGGGAATATGGAACAGGGTATATGCGTAAAAGTGTCTGTCCCAACAAAAATTTTTT 90
TTTTCCGCCTTCCCATTTTTTTTTTTTTTTTTTTGTGTGTTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTTGACCAAAAATAT 120
F S A F P F F F F F F C V F P L I S R T G Q E K L L F D Q K Y 42
AAAATTATTAAGGGCGAGAAAAAGAAAAAATCAACGAGCAAAACAGGAGAGAACACCAAAAAAGGAAATTATGCGATT 270
K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 72
AAGAAATCATTTCACATGCATCGACATGCATACGGAAGGTGAAGCAGCAGGATTGTGACGAGTGGTTTGGCACACATTCCAGGTTGCAAT 360
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 102
ATGSCGSAAGAAGAAACATACCTGCAGGAAAACATGGATTATTTGAGGCGTGGCATAATGCTGGAACACAGTGGTGCATGATGATATGTTT 430
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 132
GGAGCCTTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT 520
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 162
AACCTCAATTGCACCGGTTACGGCGGCGAGTTGAAACGGGAATTGTGAGCGTGGCGGCGAAGGCAACAAATGTTCCGGTTGTCTCGACACA 610
N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 192
CCTGCGGGTTGGTGCGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTTGCAATGCGAGTATTATCAATGTACCTCATTT 700
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 222
TTGTATCAGCAGGATGTGGTGGTGTGTGTTGCCAAAGCCCTATGGTGAAGTACGGGTTGATATTGCATTTGGAGGCAATTTTTTCGCCATT 790
L Y Q Q D V V V V L P K P Y G E V R V D I A F G S N F F A I 252
GTTCGCGCGGAGCASTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGCAGGAGAACTTCTGCGTACTGAAATCAAT 880
V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 282
CGCASTGTGAAGGTTTCAGCACCTCAGCTGCCCATATTAACTGTGGACTGTGTGAGATATACGGTCCGCCAACGAACCCGGAGGCA 970
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 312
AACACAAAGACGTTGTGATATTGGCAATCGCCAGGCGGATCCCTCTCCATGTGGGACAGGCACCAGCGCCCAAGATGGCAACACTTTAT 1060
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 342
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTTGTCCAGGGCAGGAGTACTTGGGAGGAGGCA 1150
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 372
ATACGCGGGGTGAAGGTGCCGGTGACCAAGATGCCGAGGAAGGATGCTCGTTGTAAACGGCAGAAATTAAGGAAAGGCTTTTATCATG 1240
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 402
GTTTTACACCATGCTGTTTGACCCCAACGGATCCGTTTAAGAACGATTACATTAAGCAAGTAGATCTGTTAGACACAGAACTATT 1330
G F N T M L F D P T D P F K N S F T L K Q 423
GGGGAACAGTGCAGACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTTATTTTTTATTAGTGCATT 1420
ATTATTAATTTTTTTTTTTGTTTTGGGGTTTTCAACGGTACCGGTTGGGAGCAGGGAAGCGATAGCGGCGGACAATTTTTTGTCTTTAT 1510
TTTCATTTTCATCTTCTACCCAACCCCTTGGTTCCACCGGTGCGGGCGGGTCTTGTGGTGGAGGAGTCTTAATCCCGCACCTCGG 1600
AGGAATAAACATATTTCAATTTTCATATCTTGAATCAAAAGGCAT 1651

Polyadenylation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(d) Nucleotide sequence and peptide sequence TcPA45

SEQ ID NO.8

TTTTCCGCGCTTCCCATTTTTTTTTTTTTTTTGTGTGTTTCCCTTGATCTCTCGAACAGGGCAGGAAAAGCTTCTGTTTGACCAAAATAT 130
F S A F P F F F F F C V F P L I S R T G Q E K L L F D Q K Y 140
AAAATTATTAAGGGCGAGAAAAAGAAAAAATCAACGAGCAAAACAGGAGAGAACACCAACAAAAAAGGGAAATTATGCGATTT 150
K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 160
ARGAAATCATTACATGCATCGACATGCATACGGAAGGTGAAGCAGCAGGATTGTGACGAGTGGTTTGCCACACATTCCAGGTTTCAAT 170
K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 180
ATGGCGGAGAAGAAAGCATACCTGCAGGAAAACATGGATTATTTGAGCGTGGCATAATGCTGGAACACAGTGGTCATGATGATATGTTT 190
M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D C M F 200
GGAGCCTTTTATTTGACCCTATTGAAGAAGGCGCTGACTTGGGCGATGGTATTCATGGATACCGGTGGCTATTTAAATATGTTGGACAT 210
G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 220
AACTCAATTGCAGCGGTTACGGCGGCGAGTTGAAACGGGAATTGTGAGCGTGCCGGGGAAGGCAACAAATGTTCCGTTTGTCTGGACCA 230
N S I A A V T A A V E T G I V S V P A K A T N V P V V L D T 240
CCTGGCGGGTGGTGGCGGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAAGTGTCAATGGAGTATTATCAATGTACCCCTCATTT 250
P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 260
TTTATCAAGGATGTGGTGTGTGTGTTGCCAAAGCCCTATGTTGAAGTACGGTGGTATTTGCATTTGGAGGCAATTTTTTGGCCATT 270
L Y Q Q D V V V V L P K P Y G E V R V D I A F G G N F F A I 280
GTTCGCGCGGAGCAGTTGGGAATTGATATCTCGTTTCAAAACCTCTCCAGGCTGCAGGAGGCGAGGAGAACTTCTGCCTACTGAAATCAAT 290
V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 300
CGCAGTGTGAAGGTTGAGCAGCCCTCAGCTGCCCATATTAACACTGTGGACTGTGTTGAGATATACGCTCCGCGCAACGAACCCGGAGGCA 310
R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 320
AACTACAAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCAACGCGCAAGATGGCAACACTTTAT 330
N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 340
GCCAAAGGCCAGCTTCGCATCGGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTTGGGGAGGAGGCA 350
A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E F 360
ATACCGGGGGTGAAGGTGCCGTTGACCAAGATGCCGAGGAAAGGATGCTCTTTTAAACGGCAGAAATTACTGGAAGGGCTTTTATCATG 370
I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 380
GGTTTCAACACCATGCTGTTTGAACCAACGATCCGTTTAAAGAACGATTACATTTAAAGCAGTAGATCTGTTAGAGCACAGAACTATT 390
G F N T M L F D P T D P F K N G F T L K Q 400
GGGGAACAGCTTCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATGTTTTTTTTTTATTTTTTTATTTTTTATTAGTGCATT 410
ATTATTAATTTTTTTTTTTGTTTTGGGGTTTCAACGGTACCGCGTTTGGAGCAGGGAAAGCATAGCGGCGCGGACAATTTTTTGTCTTTAT 420
TTTCATTTTTCATCTTCTACCAACCCCTTGGTTCCACCGGTGGGGCGGGTGTGTGGGTGGAGGAGTCTTAAATCCCGCACCTGG 430
AGGAATAAACATATTTCAATTTTCATATCTTGAATCAAAAGGCAT 440

Polyadenylation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(a) Nucleotide sequence and peptide sequence TcPA45

SEQ ID NO.9

F S A F P F F F F F C V F P L I S R T G Q E K L L F D Q K Y 270
 AAAATTATTAAGGGCGAGAAAAAGAAAAAATCAACGAGCAACAGGAGAGAACACCAACAAAAAGSGAAATTATGCGATT 360
 K I I K G E K K E K K K N Q R A N R R E H Q Q K R E I M R F 102
 AAGAAATCATTACATGCATCGACATGCATACGGAAGGTGAAGCAGCAGCGATTGTGACGASTGGTTTGCCACACATTCCAGSTTCGAAT 430
 K K S F T C I D M H T E G E A A R I V T S G L P H I P G S N 132
 ATGGCGGAGAAAGCATACCTGCAGGAAACATGGATTATTTGAGGCGTGCCATAATGCTGGAACACGTGGTCATGATGATATGTTT 520
 M A E K K A Y L Q E N M D Y L R R G I M L E P R G H D D M F 162
 GGAGCCTTTTTATTTGACCTATTGAAGAAGGCGCTGACTTGGGCATGGTATTCATGGATACCGGTGGCTATTTAAATATGTGTGGACAT 610
 G A F L F D P I E E G A D L G M V F M D T G G Y L N M C G H 182
 AACTCAATTGCAGCGTTACGGCGGCGAGTTGAAACGGGAATTGTGAGCGTGCCGGCGAAGGCAACAAATGTTCCGGTTGTCTCTGGACACA 700
 N S I A A V T A A V E T G I V S V P A K A T N V P Y V L D T 222
 CCGTGGGCGTTGGTGGCGGTACGGCACACCTTCAGAGTGGTACTGAGAGTGAGGTGTCAAATGCGAGTATTATCAATGTACCCCTCATT 790
 P A G L V R G T A H L Q S G T E S E V S N A S I I N V P S F 252
 TTGTATCAGCAGGATGTGGTGGTTGTGTGGCCAAAGCCCTATGGTGAATACGGGTGATATTGCATTGGAGGCAATTTTTTCGCCATT 880
 L Y Q Q D V V V L P K P Y G E V R V D I A F G S N F F A I 282
 GTTCCCGCGGAGCAGTTGGGAATTGATATCTCCGTTCAAAACCTCTCCAGGCTGCAGGAGGAGGAGAACTTCTGCTATGAAATCAAT 970
 V P A E Q L G I D I S V Q N L S R L Q E A G E L L R T E I N 312
 CGCAGTGTGAAGGTTACAGCACCTCAGCTGCCCCATATTAACACTGTGGACTGTGTGAGATATACGGTCCGCCAACGAACCCGGAGGCA 1060
 R S V K V Q H P Q L P H I N T V D C V E I Y G P P T N P E A 342
 AACTACAGAACGTTGTGATATTTGGCAATCGCCAGGCGGATCGCTCTCCATGTGGGACAGGCACCGCCCAAGATGGCAACACTTTAT 1150
 N Y K N V V I F G N R Q A D R S P C G T G T S A K M A T L Y 372
 GCCAAAGGCCAGCTTCGCATCGAGAGACTTTTGTGTACGAGAGCATACTCGGCTCACTCTTCCAGGGCAGGGTACTTGGGGAGGAGCGA 1240
 A K G Q L R I G E T F V Y E S I L G S L F Q G R V L G E E R 402
 ATACCGGGGGTGAAGGTGCCGGTGACCAAGATGCCGAGGAAGGGATGCTCGTTGTAAACGGCAGAAATTAAGTGAAGGGCTTTTATCATG 1330
 I P G V K V P V T K D A E E G M L V V T A E I T G K A F I M 423
 GGTTCACACCATGTCTTTGACCAACGGATCCGTTTAAAGACGGATTACATTAAAGCAATAGATCTGTAGAGCACAGAACTATT 1420
 G F N T M L F D P T D P F K N G F T L K Q 1510
 GGGGAACACGTGCGAACAGGTGCTGCTACGTGAAGGGTATTGAATGAATCGTTTTTTTTTATTTTTTATTTTTTATTAGTGCATT 1600
 ATTATTAATTTTTTTTTTTGTTTTGGGTTTTCAACGGTACCGCGTTGGAGCAGGGAAGCGATAGCGGCCGACAAATTTTTTCTTTTTAT 1651
 TTTTCATTTTCATCTTCCTACCCAAACCCCTTGGTTCCACCGGTCCGCGCGGGTCTTGTGGGTGGAGGASTCTAAATCCCGCACCTCGG
 AGGAATAAACATATTTCAATTTCAATCTTGAATCAAAAGGCAT

Polyadenilation site

Obs : Underlined the sequenced peptides used to deduce degenerated primers for cloning

(b) Nucleotide sequence and peptide sequence TcPA45

SEQUENCE ID NO:10

Signal peptide

1
M R K S V C P K Q K F F

Nucleotide sequence of signal sequence TcPA45

SEQUENCE ID NO:11

ATGCGATT <

X 2 7

242

222

423

22

320

242

623

212

703

222

730

22

453

112

כדן

25

943

042

د

77

245

102

553

122

229

229

— 9 —

123

552

SEQ ID NO : 12

5'TTICCRAADATIACIACGTT 3'

SEQ ID NO : 13

5' ATHGCITTYGGIGGIAAYTTT 3'

SEQ ID NO : 14

5' TTICCRAADATIACIACGTT 3'

SEQ ID NO : 15

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'

SEQ ID NO : 16

5' CTGAGCTCGACCAGATCTATCTGC 3'

SEQ ID NO : 17

```

1 ccttttttctt tttaaaaaca aaaaaaatct cgggggggaat atggaacagg gatatgctg
61 aaaagtgtct gtcccaaaaca aaaatttttt tttcccgctt tcccaatttt tttttttttt
121 tgtgtgtttt ccttgatctt tcgaacaggg caggaaaaagc ttctgtttga ccaaaaaata
181 aaaattatta agggcgagaa aaaaagaaaag aaaaaaaatc aacgagcaaa caggagagaa
241 caccaacaaa aaagggaat tatgcgattt aagaaatcat tcacatgcat cgacatgcat
301 acggaagggtg aagcagcacg gattgtgacg agtgggtttg caccacattcc aggttcgaa
361 atggcgagaa agaaaacata cctgcaggaa aacatggatt attcgaggcg tggcataatg
421 ctggaaccac gtggtcatga tgatatgttt ggagcctttt tatttcgaccc tattcgaaqa
481 ggcgctgact tgggcacgtt attcatggat accgggtggc atttaaaata gtgtggaat
541 aactcaattg cagcgggtac ggccggcagt gaaacgggaa ttgtgagcgt gccggcgaa
601 gcaacaaatg tcccggttgt cctggacaca cctgcggggt tgggtgcggg tacggcacac
661 cttcagagtg gtactgagag tgagggtgtca aatgcgagta ttatcaatgt accctcaatt
721 ttgtatcagc aggatgtggt ggttgtgttg ccaaaagccc atggtgaaat accgggttga
781 attgcatttg gaggcatttt tttcgccatt gtcccgctgg agcagttggg aattgataac
841 tccgttcaaa accctctccag gctgcaggag gcaggagaa cctctgctac tgaattcaat
901 cgcagtgtga aggttcagca ccttcagctg ccccatatta acactgttga ctgtgttga
961 atatacgttc cgccaacgaa cccggaggca aactacaaga accgtgtgat atttggcaat
1021 cgccaggcgg atcgtctctc atgtgggaca ggcaacagcg ccaagatggc aacacattat
1081 gccaaaggcc agcttcgcat cggagagact ttgtgtgac agagcaatcc cggctcaatc
1141 ttccagggca gggtaattgg ggaggagcga ataccggggg tgaaggctgc ggtgatacaa
1201 gatgccgagg aagggatgct cgttgttaac gcagaaattt ctggaaaggc ttttatcatg
1261 ggtttcaaca caatgctgtt tgaacctaac gatccgttta agaaaggatc caccattaaq
1321 cagttagatc ggttagagca agaaactatt ggggaacacg tgggaacacg tgtgtgacg
1381 tgaagggtat tgaattgaat gttttttttt aatttttttt tttattttta ttagtgaat
1441 attatttaaa tttttttttt ttttgggggt taaacgggtc cgggttggga gtagggaagc
1501 gatagcggcc ggacaatttt ttggtttttt tttcaatttt attttccctc caaaacccct
1561 tgggtccacc ggtcggcggt gggctcttgt ggtggaggag tcccaaatcc cggacccctg
1621 aggaataaac atattttcaat tccatatttt tgaattcaaa ggaat

```

SEQ ID NO : 18

WIHK

SEQ ID NO : 19

IVTGSLPDISG

SEQ ID NO : 20

ATNVPVVLDTPAGLVR

SEQ ID NO : 21

VDIAFGGNF

SEQ ID NO : 22

NVVIFGNR

SEQ ID NO : 23

MATLYAK

SEQ ID NO : 24

5' TCCGTATCCATGTCGATGC 3'

SEQ ID NO : 25

5' TATTATTGATACAGTTTCTG 3'

SEQ ID NO : 26

5' CTCTCCCATGGGGCAGGAAAAGCTTCTG 3'